MAY 2021

UNDERSTANDING NASS SURVEY DESIGNS

Dan Lofthus
MN State Statistician, USDA NASS

The June Acreage Report is the result of multiple survey projects that work together to represent all farm operations in the country. This is quite a statement to make. Farm operations start up, go out of business, change size, add or drop crops and livestock enterprises. Some add partners, some partners leave. The ag production sector is dynamic and constantly changing.

With that in mind, how can NASS claim to represent all farms in the country and account for all of these possible changes in farming and ranching? There are a few survey design rules in place to do just that.

First, **survey reference dates** are important. The NASS acreage, grain stocks, and livestock inventory reports all have specific reference dates. The reference date for the June Acreage Report, June Grain Stocks Report and the June Hog and Pig Report is June 1, 2021.

By collecting data as of June 1, 2021, you can "freeze" this dynamic ag industry, slow it down, and have confidence in measuring statistics even though things are constantly changing. For example, on that one particular day, an acre of land is most likely utilized in one particular way. It's either planted or not planted, it's corn or it's soybean, and so on. A hog is either owned by the operation on June 1 or has been sold and is no longer there on June 1. Think of it as taking an agricultural selfie on that particular day.

Second, **multiple views of agriculture** are used at the same time. Said another way, multiple survey sampling frames are used. A sampling frame is a method of organizing the subject matter (in this case agriculture) in a way that allows you to select a small group to represent the whole. NASS uses two different sampling frames for most crop and livestock surveys.

Providing Timely, Accurate and Useful Statistics In Service to U.S. Agriculture

List Sampling Frame: The first method is organizing a list of farm operators. Over time and with a lot of effort, previous surveys, Census of Agriculture returns, farm program participation and internet searches are used to build and maintain the NASS list of farm operators. This list of operators is constantly being updated as operations change and it's the primary source of farmers for NASS survey projects.

A list frame is a ready-made source of selecting farm operators for a survey, a big advantage. It's also pretty easy and inexpensive to group like operations together and randomly select a sample from that list. You can also efficiently sample for small acreage commodities or types of agriculture that are difficult to find but still important in that state. The list, if built well, also has standard ways to contact those folks, like phone numbers, mailing addresses, etc.

However, there is a disadvantage to using a list frame too. It's virtually impossible to guarantee that every farmer is on that list. In fact, we know it's not totally complete. Although we work on it constantly, there's no way to keep up with every change when farmers move, retire, change their phone number, add a cell number, or change their operation. There will always be an element of incompleteness to that list frame.

Area Sampling Frame: NASS corrects for that inherent incompleteness of the farm operator list frame by looking at agriculture in another way. Production agriculture is tied to land. Farm fields are obviously tied to land, but hog buildings, feedlots and grain bins are tied to land too. Somehow, some way, all production ag activity ties back to land.

NASS looks at agriculture based on land area by using an area sampling frame. The area sampling frame includes every single acre in the U.S. When you think about it, all agricultural activity tied to land is included in that frame simply because every acre is included. It's complete. That's the area frame's biggest advantage.

However, there are disadvantages to an area frame. The largest being cost and time required to maintain and implement this data collection in each state. It's also less efficient when collecting data for hard to find crops or confinement livestock.

To understand the area frame better, let's look at how it's built and sampled (Figure 1).

June Ag Production Survey (APS)
May 29 - June 15, 2021

Purpose:
- Update 2021 supply data used in the major crop supply and demand balance sheets.
- Provide the first look at actual planted acres of major crops in 2021.

MN Farms Sampled: 1 farm out of every 28 farms.

Reporting Options: Mail, Phone, On-line.

List Frame Survey:
  Advantages: Farm operators are known, efficient for large and small acreage crops and livestock.
  Disadvantage: It's not 100% complete.

June Area Survey
May 29 - June 15, 2021

Purpose:
- Update 2021 supply data used in the major crop supply and demand balance sheets.
- Provide the first look at actual planted acres of major crops in 2021.

MN Segments Sampled: 364

Reporting Options: Phone.

Area Frame Survey:
  Advantages: 100% complete because every acre in the state is eligible to be sampled.
  Disadvantages: Costly and time consuming, inefficient for small acreage commodities, livestock.

Step 1: Geospatial Information Systems (GIS) are used to group land into categories, or strata, based on cultivation. For example, land that is 75%+ cultivated, 50-74% cultivated and so on. This even includes land covered by towns and cities, which of course are very low in cultivation. Non-agricultural land, like lakes for example, are also identified.

Step 2: Once all land is grouped, or stratified, a sample of primary units are randomly selected. Since NASS is interested in farms and agricultural data like crop acreage and livestock inventory, highly cultivated units are included more often than units that are mostly urban.

Better Response = Better Data = Better Decisions

Step 3: Analysts, using GIS tools, work within those selected primary units to determine permanent boundaries like roads or rivers to sub-divide each selected 6-8 square mile unit into approximately 1-square mile segments of land. One of these segments, around 640 acres in size, is selected to be included in the current year's area frame survey. Since a lot of work had been done to the other segments in the unit too, they are staged to rotate into an area frame sample in future years.

There are 364 segments selected in Minnesota and over 9,000 selected across the U.S. Data are collected for each segment by screening every acre for agricultural activity. When a farmer or rancher operates land inside the boundaries of the segment, they are interviewed and asked to account for the acres they operate.

If someone lives inside the segment, they are screened to see if they have agricultural activity. By the time we are done, every acre inside the segment boundary has been accounted for and described as either agricultural or non-agricultural. Land operated, crop area, grain storage, cattle inventory, hog inventory, sheep inventory, ag labor and value of sales are collected for all agricultural operators with land inside the segment boundary.

Multiple Frame Approach to Surveys: NASS uses both types of sampling frames to maximize the advantages and minimize the disadvantages of each.

Data for samples selected from both the list and area frames are collected during the first two weeks of June and have a reference date of June 1. The June Ag Production Survey (list frame) is the primary sample and maximizes the advantages of that type of sampling frame by efficiently collecting data for large acreage crops like corn or soybean and small acreage crops like barley, oats, dry beans, canola and others. The June Area Survey (area frame) maximizes its most important benefit - being 100% complete as of June 1.

The Multiple Frame design guarantees that every farm in the U.S. had a chance of being included thanks to the June Area Survey. Plus data for a wide array of crops can be efficiently collected by the June Ag Production Survey at a much lower cost.

The area frame survey also provides a way to measure the incompleteness of the list frame. An important added benefit.

Figure 1: Example Area Frame Stratification and Sample Selection
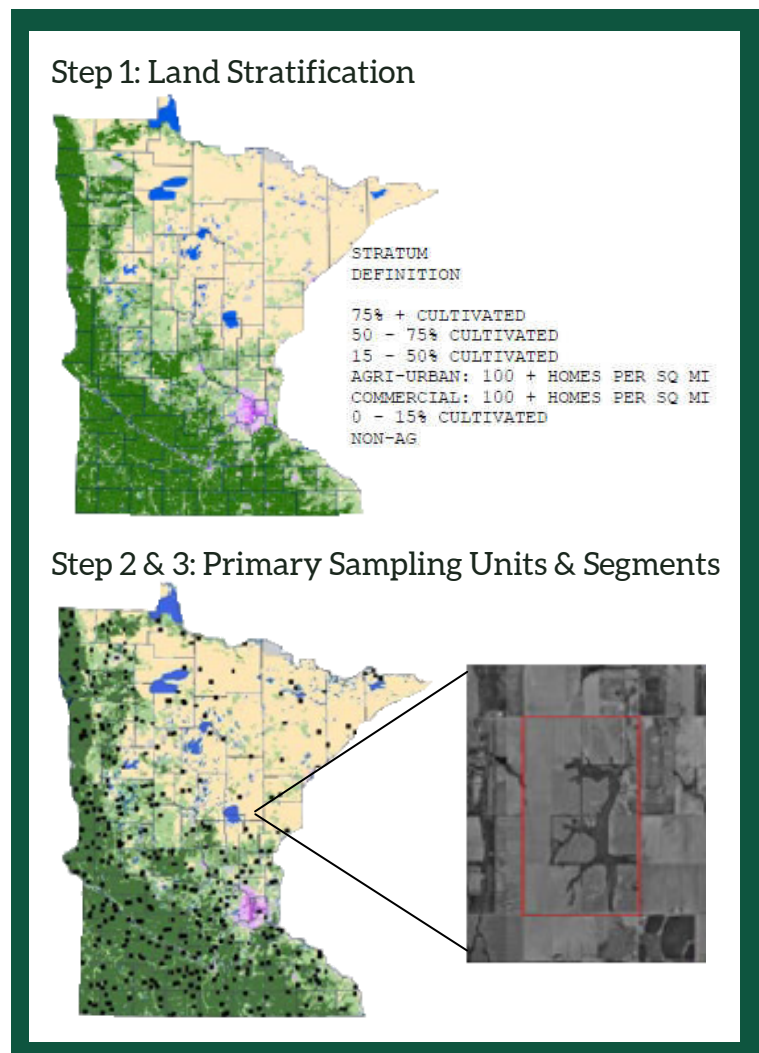
Lets use a simplified example to explain. Say there's 100,000 farms on a list frame as of June 1. By conducting an area frame survey of 250 segments we found and interviewed 2,500 farmers with agriculture as of June 1 using that unique sampling frame.

Now we have two independent lists of farm operators developed (Figure 2). By matching the operators found on the area frame to the operators on the list frame, you can measure the completeness of that list frame.

Also, since we are collecting crops, stocks, cattle, hogs, sheep, ag labor, and more, we can measure the completeness of each commodity type as well.

Using the same example, lets say we find that 125 area operators are not found on the list frame after matching is complete. Based on this example, the incompleteness of the list frame would be measured at 5% (125/2,500). Using this same method of matching, we can evaluate the coverage of the different commodities too.

Continuing the example, each June, 40 new segments would be rotated into the area sample of 200 and 40 would be rotated out. That's a 20% rotation per year so each segment is included for 5 years. You could select a completely new 200 segment sample each year but the startup cost of a brand new segment is significant. Once we invest in that start up cost, a segment is kept in the area sample for 5 years. Each segment is screened closely every year to capture changes. The matching step is redone each year when new data is collected and an updated measure of list frame completeness is calculated.

The list frame accounts for the majority of the data collected but the area frame contributes data for those operations that are not represented by the list frame. A very important function to ensure all farms had an opportunity to be included in the survey program.

We continue the multiple frame survey approach all year long by including some of those area frame non-match farms in the samples for major crop, livestock and economic surveys for the next 11 months.

The matching process also makes sure that a farm's data is represented by either the list sampling frame or the area sampling frame but not both frames. This avoids the obvious possibility of duplicating that data.

Figure 2: Multiple Frame Example Measuring Incompleteness Each Year
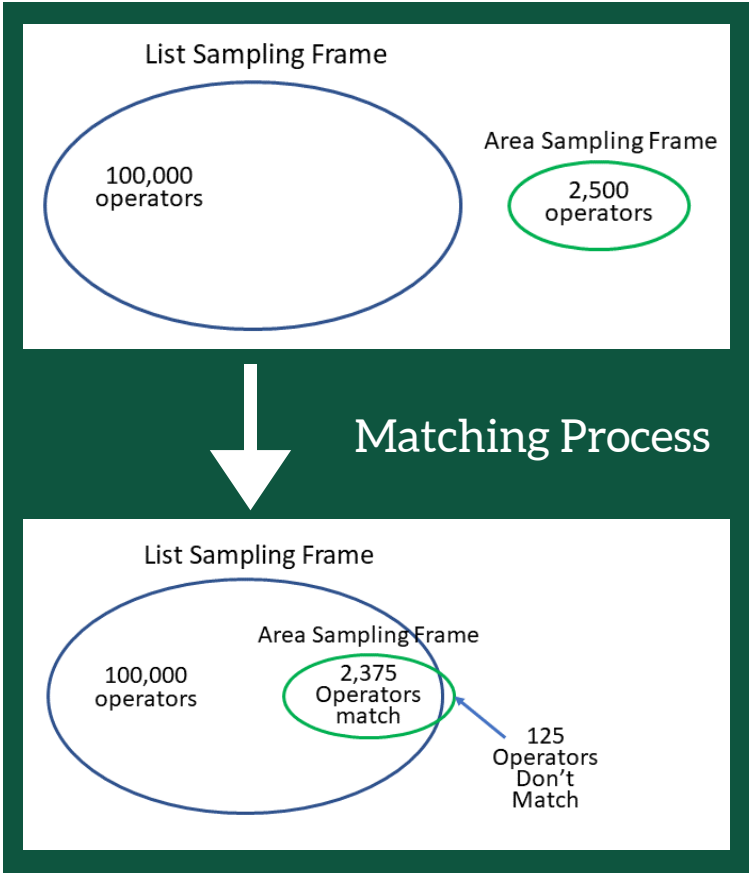
The NASS multiple frame survey design is used for 10 estimation programs including 24 major surveys during the year (Figure 3). Several include estimates for speculative commodities and market sensitive reports. So when NASS makes the statement that every farm is represented in the results of these surveys, you now know that's accomplished by using a multiple frame survey design.

Figure 3: NASS Survey Programs using the Multiple Frame Design

Survey Program	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May
Crop Surveys	X			X			X			X		
On Farm Grain Stocks	X			X			X			X		
Cattle Inventory		X						X				
Hog Inventory	X			X			X			X		
Sheep Inventory								X				
Goat Inventory								X				
Farm Numbers	X						X					
Land in Farms	X						X					
Ag Labor					X						X	
Ag Resource Management Studies (ARMS)					X				X			